SYMMETRIC APPROXIMATIONS OF PSEUDO-BOOLEAN FUNCTIONS

JEAN-LUC MARICHAL AND PIERRE MATHONET

ABSTRACT. We consider the approximation problem of a pseudo-Boolean function by a symmetric pseudo-Boolean function in the sense of weighted least squares. We give explicit expressions for the approximation and provide interpretations and properties of its *L*-statistic representation. We also discuss applications of these expressions in cooperative game theory and engineering reliability.

1. Introduction

Hammer and Holzman [4] investigated the problem of approximating a pseudo-Boolean function $f: \{0,1\}^n \to \mathbb{R}$ by another pseudo-Boolean function of smaller degree in the sense of standard least squares, where the degree of a pseudo-Boolean function is that of its unique multilinear polynomial representation. As an application, they showed that the Banzhaf power index, an important concept in cooperative game theory, appears as the coefficients of the linear terms in the solution of the approximation problem by functions of degree at most one.

Motivated by this work, in this paper we consider the problem of approximating a pseudo-Boolean function by a symmetric one (i.e., invariant under permutation of its variables). That is, we consider the closest symmetric pseudo-Boolean function to a given pseudo-Boolean function in the sense of least squares. To be general enough, we actually consider arbitrarily weighted least squares. We give explicit expressions for the approximations as symmetric multilinear polynomials and shifted L-statistic functions and we discuss a few properties of these approximations (Section 2). We then investigate the shifted L-statistic representation of the approximations and interpret their coefficients as a way to measure the influence of the kth largest variable (for arbitrary $k \leq n$) on the given pseudo-Boolean function (Section 3). Finally, we discuss two applications of the approximation problem: defining a new concept of influence of players in cooperative game theory and interpreting system signatures used in engineering reliability (Section 4).

We employ the following notation throughout the paper. We denote by \mathbb{B} the 2-element set $\{0,1\}$. For any $S \subseteq [n] = \{1,\ldots,n\}$, we denote by $\mathbf{1}_S$ the characteristic vector of S in $\{0,1\}^n$ (with the particular case $\mathbf{0} = \mathbf{1}_{\varnothing}$).

Recall that if the \mathbb{B} -valued variables x_1, \ldots, x_n are rearranged in ascending order of magnitude $x_{(1)} \leqslant \cdots \leqslant x_{(n)}$, then $x_{(k)}$ is called the *kth order statistic* and the function $os_k : \mathbb{B}^n \to \mathbb{B}$, defined as $os_k(\mathbf{x}) = x_{(k)}$, is the *kth order statistic function*.

1

Date: April 15, 2010.

 $^{2010\} Mathematics\ Subject\ Classification.$ Primary 41A10, 93E24; Secondary 62G30, 90B25, 91A12.

Key words and phrases. Pseudo-Boolean function, least squares approximation, symmetric function, Lovász extension, cooperative game theory, system reliability, system signature.

We then have $os_k(\mathbf{x}) = 1$, if $\sum_{i=1}^n x_i \ge n - k + 1$, and 0, otherwise. As a matter of convenience, we also formally define $os_0 \equiv 0$ and $os_{n+1} \equiv 1$. An *L-statistic* function is a linear combination of the functions os_1, \ldots, os_n while a *shifted L-statistic* function is a linear combination of the functions os_1, \ldots, os_{n+1} .

2. Symmetric approximations

In this section we present and solve the approximation problem of pseudo-Boolean functions by symmetric pseudo-Boolean functions and we discuss a few properties of the approximations.

Through the usual identification of the elements of \mathbb{B}^n with the subsets of [n], an n-ary pseudo-Boolean function $f: \mathbb{B}^n \to \mathbb{R}$ can be equivalently described by a set function $v_f: 2^{[n]} \to \mathbb{R}$. The correspondence is given by $v_f(S) = f(\mathbf{1}_S)$ and

(1)
$$f(\mathbf{x}) = \sum_{S \subseteq [n]} v_f(S) \prod_{i \in S} x_i \prod_{i \in [n] \setminus S} (1 - x_i).$$

Equation (1) shows that any n-ary pseudo-Boolean function f can always be represented by a multilinear polynomial of degree at most n (see [3]), which can be further simplified into

(2)
$$f(\mathbf{x}) = \sum_{S \subseteq [n]} a_f(S) \prod_{i \in S} x_i,$$

where the set function $a_f: 2^{[n]} \to \mathbb{R}$, called the Möbius transform of v_f , is defined by

$$a_f(S) = \sum_{T \subset S} (-1)^{|S| - |T|} v_f(T).$$

The following proposition gives two equivalent conditions for a pseudo-Boolean function to be symmetric, including a description as a shifted L-statistic function. A set function $v\colon 2^{[n]}\to \mathbb{R}$ is said to be $\operatorname{cardinality-based}$ if v(S)=v(T) for every $S,T\subseteq [n]$ such that |S|=|T|. Equivalently, there exists a unique function $\overline{v}\colon \{0,1,\ldots,n\}\to \mathbb{R}$ such that $v(S)=\overline{v}(|S|)$.

Proposition 1. If f is a pseudo-Boolean function then the following assertions are equivalent.

- (i) f is symmetric.
- (ii) v_f is cardinality-based.
- ${\rm (iii)}\ \ f\ \ is\ a\ shifted\ L\text{-}statistic\ function.}$

Proof. $(iii) \Rightarrow (i) \Leftrightarrow (ii)$ Trivial.

 $(ii) \Rightarrow (iii)$ If v_f is cardinality-based, then using (1) we obtain

$$f(\mathbf{x}) = \sum_{s=0}^{n} \overline{v}_f(s) \sum_{\substack{S \subseteq [n] \\ |S|}} \prod_{i \in S} x_i \prod_{i \in [n] \setminus S} (1 - x_i) = \sum_{s=0}^{n} \overline{v}_f(s) \left(x_{(n-s+1)} - x_{(n-s)} \right),$$

which completes the proof.

Denote by $F(\mathbb{B}^n)$ the set of *n*-ary pseudo-Boolean functions and by $F_S(\mathbb{B}^n)$ the subset of symmetric *n*-ary pseudo-Boolean functions. Since $F_S(\mathbb{B}^n)$ is precisely the set of *L*-statistic functions (see Proposition 1), it is spanned by the linearly independent set $B = \{os_1, \ldots, os_n, os_{n+1}\}$ and thus is a linear subspace of $F(\mathbb{B}^n)$ of dimension n+1.

Given a weight function $w \colon \mathbb{B}^n \to]0, \infty[$ and a function $f \in F(\mathbb{B}^n)$, we define the best symmetric approximation of f as the unique function $f_L \in F_S(\mathbb{B}^n)$ that minimizes the weighted squared distance

$$||f - g||^2 = \sum_{\mathbf{x} \in \mathbb{R}^n} w(\mathbf{x}) (f(\mathbf{x}) - g(\mathbf{x}))^2$$

among all symmetric functions $g \in F_S(\mathbb{B}^n)$. Note that $\|\cdot\|$ is the norm associated with the inner product $\langle f, g \rangle = \sum_{\mathbf{x} \in \mathbb{B}^n} w(\mathbf{x}) f(\mathbf{x}) g(\mathbf{x})$.

We can assume w.l.o.g. that the weights w(S) are multiplicatively normalized so that $\sum_{\mathbf{x} \in \mathbb{B}^n} w(\mathbf{x}) = 1$. We then immediately see that the weights define a probability distribution over \mathbb{B}^n .

Clearly, the solution of this approximation problem exists and is uniquely determined by the orthogonal projection of f onto $F_S(\mathbb{B}^n)$. We then write $f_L = A(f)$.

The next theorem gives an explicit expression for A(f). For every $f \in F(\mathbb{B}^n)$, define $\overline{v}_f : \{0, 1, ..., n\} \to \mathbb{R}$ as $\overline{v}_f(s) = E(f(\mathbf{x}) \mid |\mathbf{x}| = s)$. In the special case when the weight function w is symmetric, we obtain

$$\overline{v}_f(s) = \frac{1}{\binom{n}{s}} \sum_{|\mathbf{x}|=s} f(\mathbf{x}).$$

Theorem 2. The best symmetric approximation of $f \in F(\mathbb{B}^n)$ is given by

(3)
$$A(f) = \sum_{j=1}^{n+1} c_j os_j,$$

where $c_i = \overline{v}_f(n-j+1) - \overline{v}_f(n-j)$ for every $j \in [n]$ and $c_{n+1} = \overline{v}_f(0) = f(\mathbf{0})$.

Proof. The projection A(f) is characterized by the conditions

$$\langle f - A(f), os_i \rangle = 0 \qquad (i \in [n+1]),$$

that is,

(5)
$$\sum_{|\mathbf{x}| \geqslant n-i+1} w(\mathbf{x}) (f(\mathbf{x}) - A(f)(\mathbf{x})) = 0 \qquad (i \in [n+1]).$$

We observe that conditions (5) remain equivalent if we replace the inequality $|\mathbf{x}| \ge n - i + 1$ with the equality. Using (3), we then obtain

(6)
$$\left(\sum_{j=i}^{n+1} c_j\right) \left(\sum_{|\mathbf{x}|=n-i+1} w(\mathbf{x})\right) = \sum_{|\mathbf{x}|=n-i+1} w(\mathbf{x}) f(\mathbf{x}) \qquad (i \in [n+1])$$

We finally obtain the result by subtracting equation i+1 from equation i.

The next corollary yields alternative expressions for A(f) as a shifted L-statistic function and symmetric multilinear polynomials.

Corollary 3. The best symmetric approximation of $f \in F(\mathbb{B}^n)$ is given by

(7)
$$A(f) = \langle f, 1 \rangle + \sum_{j=1}^{n} c_j (\operatorname{os}_j - \langle \operatorname{os}_j, 1 \rangle),$$

¹Note that this notation is consistent with that introduced for cardinality-based set functions.

where $c_j = \overline{v}_f(n-j+1) - \overline{v}_f(n-j)$ for every $j \in [n]$. Also,

(8)
$$A(f)(\mathbf{x}) = \sum_{S \subseteq [n]} \overline{v}_f(|S|) \prod_{i \in S} x_i \prod_{i \in [n] \setminus S} (1 - x_i) = \sum_{S \subseteq [n]} \overline{a}_{A(f)}(|S|) \prod_{i \in S} x_i,$$

where $\overline{a}_{A(f)}(s) = \Delta_k^s \, \overline{v}_f(k)|_{k=0}$.

Proof. Equation (7) follows from (3) and (4) (for i = n + 1). To establish (8), by (1) and (2) we only need to prove that $\overline{v}_{A(f)}(s) = \overline{v}_f(s)$ and $\overline{a}_{A(f)}(s) = \Delta_k^s \overline{v}_f(k)|_{k=0}$. Let $S \subseteq [n]$ such that |S| = s. Then, using (3) and (6), we obtain

$$\overline{v}_{A(f)}(s) = A(f)(\mathbf{1}_S) = \sum_{j=n-s+1}^{n+1} c_j = E(f(\mathbf{x}) \mid |\mathbf{x}| = s) = \overline{v}_f(s)$$

and hence, since $a_{A(f)}$ is cardinality-based,

$$\overline{a}_{A(f)}(s) = a_{A(f)}(S) = \sum_{T \subseteq S} (-1)^{s-|T|} \overline{v}_{A(f)}(|T|) = \sum_{t=0}^{s} {s \choose t} (-1)^{s-t} \overline{v}_{A(f)}(t)$$
$$= \Delta_k^s \overline{v}_f(k)|_{k=0},$$

which completes the proof.

We now examine the effect of a permutation of the variables of f on the symmetric approximation A(f). Let S_n denote the symmetric group on [n]. A permutation $\pi \in S_n$ acts on a pseudo-Boolean function $f \in F(\mathbb{B}^n)$ by $\pi(f)(x_1, \ldots, x_n) = f(x_{\pi(1)}, \ldots, x_{\pi(n)})$. A permutation $\pi \in S_n$ is said to be a *symmetry* of $f \in F(\mathbb{B}^n)$ if $\pi(f) = f$.

Proposition 4. If $\pi \in S_n$ is a symmetry of the weight function w, then for every $f \in F(\mathbb{B}^n)$ we have $A(\pi(f)) = A(f)$ and $\|\pi(f) - A(f)\| = \|f - A(f)\|$.

Proof. If π is a symmetry of w, then clearly it is an isometry of $F(\mathbb{B}^n)$, that is, $\langle \pi(f), \pi(g) \rangle = \langle f, g \rangle$. Now, if $g \in F_S(\mathbb{B}^n)$, then by (4), we have $\langle \pi(f), g \rangle = \langle \pi(f), \pi(g) \rangle = \langle f, g \rangle = \langle A(f), g \rangle$, which shows that $A(\pi(f)) = A(f)$. Using a similar argument, we obtain

$$\|\pi(f) - A(f)\|^2 = \langle \pi(f) - A(f), \pi(f) - A(f) \rangle = \langle f - A(f), f - A(f) \rangle$$

= $\|f - A(f)\|^2$,

which completes the proof.

With any pseudo-Boolean function $f \in F(\mathbb{B}^n)$, we can associate the symmetric function $\operatorname{Sym}(f) = \frac{1}{n!} \sum_{\pi \in S_n} \pi(f)$. We then have the following result.

Corollary 5. If the weight function w is symmetric, then for every $f \in F(\mathbb{B}^n)$ we have $\operatorname{Sym}(f) = A(\operatorname{Sym}(f)) = A(f)$.

Proof. The first equality follows from the symmetry of Sym(f). The second one follows from Proposition 4 and the linearity of the projector A.

Remark 1. Let $F_k(\mathbb{B}^n)$ be the class of pseudo-Boolean functions of degree less than or equal to k and let A_k be the orthogonal projector onto $F_k(\mathbb{B}^n)$ (see [2, 4]). If w is symmetric, then the best approximation of $f \in F(\mathbb{B}^n)$ by an element of $F_S(\mathbb{B}^n) \cap F_k(\mathbb{B}^n)$ is given by $A(A_k(f))$. In fact, by Corollary 5, the function $A(A_k(f))$ lies in $F_S(\mathbb{B}^n) \cap F_k(\mathbb{B}^n)$ and we have

$$f - A(A_k(f)) = (f - A_k(f)) + (A_k(f) - A(A_k(f))),$$

where $f - A_k(f)$ is orthogonal to $F_k(\mathbb{B}^n)$ and $A_k(f) - A(A_k(f))$ is orthogonal to $F_S(\mathbb{B}^n)$. Using the properties of A_k , we can also show that $A_k(A(f)) = A(A_k(f))$.

We end this section by analyzing the effect of dualization of f on the symmetric approximation A(f). The dual of a function $f \in F(\mathbb{B}^n)$ is the function $f^d \in F(\mathbb{B}^n)$ defined by $f^d(\mathbf{x}) = 1 - f(\mathbf{1}_{[n]} - \mathbf{x})$. The following result is straightforward.

Proposition 6. If the weight function w satisfies $w(\mathbf{1}_{[n]} - \mathbf{x}) = w(\mathbf{x})$, then for every $f \in F(\mathbb{B}^n)$ we have $\overline{v}_{f^d}(s) = 1 - \overline{v}_f(n-s)$.

3. Influence of the kth largest variable

Following Hammer and Holzman's approach [4], to measure the global influence of the kth largest variable $x_{(k)}$ on an arbitrary pseudo-Boolean function $f \in F(\mathbb{B}^n)$, it is natural to define an index $I: F(\mathbb{B}^n) \times [n] \to \mathbb{R}$ as $I(f,k) = c_k$, where c_k is defined in Theorem 2.

Definition 7. Let $I: F(\mathbb{B}^n) \times [n] \to \mathbb{R}$ be defined as $I(f,k) = \overline{v}_f(n-k+1)$ $\overline{v}_f(n-k)$.

Thus we have defined an influence index from an elementary approximation (projection) problem. Conversely, the following result shows that A(f) is the unique function of $F_S(\mathbb{B}^n)$ that preserves the average value and the influence index of f.

Proposition 8. A function $g \in F_S(\mathbb{B}^n)$ is the best symmetric approximation of $f \in F(\mathbb{B}^n)$ if and only if $\langle f, 1 \rangle = \langle g, 1 \rangle$ and I(f, k) = I(g, k) for all $k \in [n]$.

Proof. The necessity is trivial (use Eq. (4) for i = n + 1). To prove the sufficiency, observe that any $g \in F_S(\mathbb{B}^n)$ satisfying the assumptions of the proposition is of the form

$$g = g(\mathbf{0}) + \sum_{j=1}^{n} I(g,j) \operatorname{os}_{j} = g(\mathbf{0}) + \sum_{j=1}^{n} I(f,j) \operatorname{os}_{j}.$$
 We then have $g(\mathbf{0}) + \sum_{j=1}^{n} I(f,j) \langle \operatorname{os}_{j}, 1 \rangle = \langle g, 1 \rangle = \langle f, 1 \rangle$. Using (7), we finally

obtain q = A(f).

We now give a few properties of the index I(f, k).

First, we see from Proposition 4 that $I(\pi(f), k) = I(f, k)$ whenever π is a symmetry of w. Similarly, we see from Corollary 5 that $I(f,k) = I(\operatorname{Sym}(f),k)$ whenever w is symmetric.

We also see from Proposition 6 that, if w satisfies $w(\mathbf{1}_{[n]} - \mathbf{x}) = w(\mathbf{x})$, then $I(f^d, k) = I(f, n - k + 1).$

We also have the following immediate property.

Proposition 9. For every fixed $k \in [n]$, the mapping $f \mapsto I(f, k)$ is linear.

It is a well-known fact of linear algebra that a linear map on a finite dimensional inner product space can be expressed as an inner product with a fixed vector. The next proposition gives the explicit form of such a vector for $I(\cdot, k)$. To this extent, for every $k \in [n]$ we introduce the function $g_k : \mathbb{B}^n \to \mathbb{R}$ as $g_k(\mathbf{x}) = \Delta_k(d_k \Delta_k os_{k-1})$, where $d_k = -1/\sum_{|x|=n-k+1} w(\mathbf{x})$.

Proposition 10. For every $f \in F(\mathbb{B}^n)$ and every $k \in [n]$, we have $I(f,k) = \langle f, g_k \rangle$. *Proof.* We have $d_{k+1}\langle f, \Delta_k os_k \rangle = d_{k+1} \sum_{|\mathbf{x}|=n-k} w(\mathbf{x}) f(\mathbf{x}) = -\overline{v}_f(n-k)$, which leads immediately to the result.

Proposition 10 shows that the index I(f,k) is the covariance of the random variables f and g_k . Indeed, we have $I(f,k) = E(f\,g_k) = \cot(f,g_k) + E(f)\,E(g_k)$, where $E(g_k) = \langle 1,g_k \rangle = I(1,k) = 0$. From the usual interpretation of the concept of covariance, we see that an element $\mathbf{x} \in \mathbb{B}^n$ has a positive contribution on I(f,k) whenever the values of $f(\mathbf{x}) - E(f)$ and $g_k(\mathbf{x}) - E(g_k) = g_k(\mathbf{x})$ have the same sign. Note that $g_k(\mathbf{x})$ is positive whenever $x_{(k)}$ is greater than the value $(d_{k+1}x_{(k+1)} + d_kx_{(k-1)})/(d_{k+1} + d_k)$, which lies in the range of $x_{(k)}$ when the other order statistics are fixed at \mathbf{x} .

4. Applications

We now discuss two applications of the best symmetric approximation: defining a new concept of influence of players in cooperative game theory and interpreting system signatures used in engineering reliability.

4.1. Influence of players in cooperative games. In cooperative game theory, the set [n] stands for the set of *players*, and its subsets are called *coalitions*. A *game* is a set function $v: 2^{[n]} \to \mathbb{R}$ with $v(\emptyset) = 0$. One thinks of v(S) as the worth of the coalition S. A game v is said to be *symmetric* if v is cardinality-based.

In applications, it may be useful to approximate a given game v on [n] by a simpler one, for instance by a game whose corresponding pseudo-Boolean function has a degree strictly less than n. This problem was solved in [2,4], where interpretations of the so-called Banzhaf power and interaction indexes were given. Alternatively, it may be interesting to consider the best symmetric approximation of v, that is, the closest symmetric game to v. Since v can be seen as a pseudo-Boolean function $f \in F(\mathbb{B}^n)$ that assumes the value 0 at $\mathbf{0}$, the closest symmetric game to v is precisely the game defined by A(f). We can then interpret the influence index $I(f,k) = \overline{v}(n-k+1) - \overline{v}(n-k)$ as the mean contribution of an arbitrary player to an arbitrary (n-k)-coalition.

4.2. System signatures in engineering reliability. Consider a system consisting of n components. When the components have i.i.d. lifetimes X_1, \ldots, X_n , the signature of the system is defined as the n-tuple $(s_1, \ldots, s_n) \in [0, 1]^n$ with $s_i = \Pr(T = X_{i:n})$, where T denotes the system lifetime. That is, s_i is the probability that the ith component failure causes the system to fail. (For a recent reference, see [6].) It was proved [1] that

$$s_i = \frac{1}{\binom{n}{n-i+1}} \sum_{|\mathbf{x}|=n-i+1} \phi(\mathbf{x}) - \frac{1}{\binom{n}{n-i}} \sum_{|\mathbf{x}|=n-i} \phi(\mathbf{x}),$$

where $\phi \colon \mathbb{B}^n \to \mathbb{B}$ is the structure function of the system. Thus we have $s_i = \overline{v}_{\phi}(n-i+1) - \overline{v}_{\phi}(n-i)$, assuming that the weight function w is symmetric. We then observe that s_i is precisely the index $I(\phi,i)$, which can be interpreted as the influence on the system of the component that has the ith largest lifetime.²

It follows that any expression involving the system signature can be reformulated in terms of the influence index for the structure function ϕ . For instance, as far as the system reliability is concerned, we have the following theorem (see [6, p. 27]).

²Note that this influence index is obtained from a standard least squares approximation problem of the structure function.

Theorem 11. For i.i.d. lifetimes X_1, \ldots, X_n , we have $\overline{F}(t) = \sum_{i=1}^n I(\phi, i) \overline{F}_{i:n}(t)$ for all t > 0, where $\overline{F}(t) = \Pr(T > t)$ and $\overline{F}_{i:n}(t) = \Pr(X_{i:n} > t)$.

5. Conclusion

We have defined an influence index from the weighted least squares approximation problem of pseudo-Boolean functions by symmetric ones, with an interpretation in game theory ($\S4.1$). In the non-weighted case, where the approximation reduces to symmetrization in the usual sense (Corollary 5), we have shown that the influence index coincides with the signature of a system made up of i.i.d. components ($\S4.2$).

ACKNOWLEDGMENTS

This research is supported by the internal research project F1R-MTH-PUL-09MRDO of the University of Luxembourg.

References

- [1] Ph. J. Boland. Signatures of indirect majority systems. J. Appl. Prob., 38:597-603, 2001.
- [2] M. Grabisch, J.-L. Marichal, and M. Roubens. Equivalent representations of set functions. Math. Oper. Res., 25(2):157–178, 2000.
- [3] P. Hammer and S. Rudeanu. Boolean methods in operations research and related areas. Berlin-Heidelberg-New York: Springer-Verlag, 1968.
- [4] P. Hammer and R. Holzman. Approximations of pseudo-Boolean functions; applications to game theory. Z. Oper. Res., 36(1):3–21, 1992.
- [5] J.-L. Marichal and P. Mathonet. The influence of the kth largest variable on functions over the unit hypercube. arXiv:1003.2556.
- [6] F.J. Samaniego. System signatures and their applications in engineering reliability. Int. Series in Operations Research & Management Science, 110. New York: Springer, 2007.

Mathematics Research Unit, FSTC, University of Luxembourg, 6, rue Coudenhove-Kalergi, L-1359 Luxembourg, Grand Duchy of Luxembourg.

E-mail address: jean-luc.marichal[at]uni.lu

MATHEMATICS RESEARCH UNIT, FSTC, UNIVERSITY OF LUXEMBOURG, 6, RUE COUDENHOVE-KALERGI, L-1359 LUXEMBOURG, GRAND DUCHY OF LUXEMBOURG.

E-mail address: pierre.mathonet[at]uni.lu